De omstreden kwaliteit van integriteitsonderzoek dat topambtenaren en bestuurders hun baan kost

Liza Boon
Opleiding Bestuurskunde
Faculteit der Managementwetenschappen
Radboud Universiteit Nijmegen
Maart 2017
Begeleider: Prof. dr. M. S. De Vries

Samenvatting

Integriteitsonderzoek wordt door onderzoekers met uiteenlopende achtergronden uitgevoerd. Duidelijke richtlijnen voor integriteitsonderzoek zijn er niet, maar integriteitsonderzoek kan wel grote consequenties hebben voor degene waarnaar het onderzoek wordt gedaan. Het doel van dit artikel is om in kaart te brengen in hoeverre integriteitsonderzoek voldoet aan de kwaliteitseisen van wetenschappelijk onderzoek. Om hierachter te komen, zijn er 47 rapporten geanalyseerd die tussen 2010 en 2015 zijn opgesteld en over integriteitsschendingen in het openbaar bestuur gaan. In de analyse zijn er drie fases van elkaar te onderscheiden: de fase voorafgaand aan het onderzoek, de fase van de uitvoering van het onderzoek en de fase van de afronding/verslaglegging van het onderzoek. Het blijkt dat in alle drie deze fases van integriteitsonderzoek zaken zijn die misgaan. Veel voorkomende problemen bij het doen van integriteitsonderoek is de bemoeienis van de opdrachtgever, het niet meewerken van betrokkenen en breed interpreteerbare normenkaders. In slechts 13 van de 47 rapporten zijn de getrokken conclusies houdbaar.

1 Inleiding

Bestuurders in spé, bestuurskundestudenten en alumni, vragen weleens wat de zin is van al dat onderwijs in methoden en technieken van onderzoek gedurende de studie. Zij worden immers geen onderzoekers, maar gaan het openbaar bestuur in als beleidsmedewerkers, managers of bestuurders. Dit artikel zal verduidelijken dat kennis van onderzoekmethoden ook zinvol is, omdat zij tijdens hun loopbaan kunnen worden geconfronteerd met onderzoek dat hen terecht of onterecht hun baan kan kosten. Bekende voorbeelden van topambtenaren en bestuurders die op basis van onderzoek naar hun integriteit zijn ontslagen of ontslag hebben moeten nemen betreffen de oud-burgemeester van Schiedam, de oud-gemeentesecretaris van de gemeente Tynaarlo, de oud-directeur van de Stadsbank Oost Nederland en eerder de directeur van het Centraal Orgaan Asielzoekers.

Gegeven de vergaande consequenties die zulk onderzoek kan hebben, mag worden verondersteld dat de onderzoekers zo zorgvuldig mogelijk te werk gaan en zoveel mogelijk trachten te voldoen aan de onderzoekvereisten die vanuit de wetenschap dan wel vanuit hun beroepsgroep worden gesteld. Of de inhoud van zulk onderzoek in al haar fasen inderdaad voldoet aan kwaliteitseisen, vormt de centrale vraag van dit artikel. Die vraag is relevant, omdat recentelijk een aantal betrokken bestuurders de kwaliteit van zulk onderzoek heeft betwist, soms zelfs naar de rechter is gestapt en via civiele procedures heeft getracht het betreffende onderzoeksbureau te laten betalen voor de aangerichte schade. In februari 2015 werden vragen gesteld in de Tweede Kamer over het onderzoek van een van die integriteitsbureaus (BING) en werd gevraagd om een meta-evaluatie van zulk integriteitsonderzoek. In de beantwoording van deze vragen stelde de minister daar geen belang in te stellen, gegeven het onderzoek dat al was verricht naar de vormgeving van zulk onderzoek (Zouridis & Vorm, 2013) en de handreiking die het ministerie van binnenlandse zaken voor dergelijk onderzoek al had opgesteld. Hij had ook kunnen wijzen op het rapport van Addink (2013) over de kwaliteit van het bestuurlijk-juridische toetsingskader van zulk onderzoek. Recentelijk hebben ook gemeenten protocollen ontwikkeld op basis waarvan integriteitsonderzoek dient te worden verricht (zie bijvoorbeeld de gemeenten Amsterdam, Den Haag, Waalwijk).

Die documenten geven een adequaat overzicht van de formeel te volgen procedures, de context waarin zulk onderzoek plaatsvindt, de fasering van zulk onderzoek, de variërende werkwijze van betrokken onderzoeksbureaus en de dilemma's die kunnen spelen. Ze gaan echter niet in op de inhoud van dergelijk onderzoek, of op de vraag of de conclusies die in zulke integriteitsrapporten worden getrokken altijd gebaseerd zijn op onderzoek dat de toets der kritiek kan doorstaan.

Dat laatste is wel het doel van dit artikel. Het geeft voorbeelden van methodologische missers in integriteitsonderzoek die relevant zijn voor de getrokken conclusies over de betrokken topambtenaar of bestuurder. De uitkomsten die hieronder worden beschreven moeten worden gezien als het resultaat van een beschrijvend, exploratief onderzoek dat slechts de pretentie heeft te illustreren op welke wijze integriteitsonderzoek niet voldoet aan methodologische kwaliteitseisen. Dit onderzoek vormt geen meta-evaluatie in de formele betekenis van het woord, maar geeft wel de noodzaak daarvan aan, want slechts

een minderheid van de geanalyseerde integriteitsonderzoeken voldeed methodologisch volledig aan de kwaliteitseisen.

De bestudeerde rapporten beperken zich tot interne en externe onderzoekingen tussen 2010 en 2015 in Nederland die zodanig kritisch waren naar een topambtenaar of bestuurder werkzaam in het openbaar bestuur dat die daardoor in problemen kwam. In 8 van de 47 hier geanalyseerde integriteitsonderzoeken gaat het om intern onderzoek en in twee gevallen om een samenwerking tussen interne en externe onderzoekers. In het onderhavige onderzoek werd 13 keer een hoogleraar gevraagd het onderzoek te doen, acht keer werd het onderzoek intern gedaan, zeven keer een extern organisatieadviesbureau en zes keer was het een externe instantie gespecialiseerd in integriteitsonderzoek en beleid, zoals BING, die het onderzoek deed,. Vijf keer ging het om externe forensische accountants, vier keer om juridische dienstverleners of advocaten, en twee keer om externe recherchebureaus.

Met deze beperking in de scope kan niet worden uitgesloten dat er meer integriteitsonderzoek bestaat dat wel aan de maat is. Daarover worden geen uitspraken gedaan.

Gegeven de vraagstelling of de inhoud van dergelijk onderzoek in al haar fasen voldoet aan daaraan te stellen kwaliteitseisen, geeft dit artikel allereerst een overzicht van te stellen kwaliteitseisen en van de fasen die in zulk onderzoek worden onderscheiden. Daarna worden per onderzoeksfase voorbeelden gegeven van wat inhoudelijk is te zien in die rapporten dat niet strookt met de kwaliteitseisen. Het geheel wordt afgesloten met een discussie van de uitkomsten en conclusies.

2 Kwaliteitseisen vanuit wetenschap en praktijk

In de wetenschap gelden de validiteit en betrouwbaarheid als standaardcriteria voor de kwaliteit van onderzoek. Validiteit heeft betrekking op de vraag of gemeten is wat men beoogt te meten en de betrouwbaarheid betreft de vraag of het onderzoek zodanig is uitgevoerd en gerapporteerd dat een andere onderzoeker die op dezelfde wijze te werk gaat ook tot dezelfde uitkomsten zou komen.

Nu kan worden gesteld dat integriteitsonderzoek niet zonder meer aan wetenschappelijke criteria kan voldoen gegeven de specifieke kenmerken ervan. Deze kenmerken zijn recentelijk beschreven door Zouridis en van der Vorm (2013). Zij stellen dat dergelijk onderzoek een bijzondere tak van sport is vanwege de volgende redenen: ‘Het vergt de precisie en de zorgvuldigheid van strafrechtelijk opsporingsonderzoek, maar ook de sensitiviteit om onderzoek te doen in een conflictueuze omgeving en de beheersing van uiteenlopende disciplines als arbeidsrecht en organisatiesociologie’. Het is ook problematisch vanwege de ingrijpende gevolgen die zulk onderzoek kan hebben voor betrokkenen. ‘Reputatieschade laat zich niet snel repareren als gevolg van de sterke morele connotatie van integriteit’, Dit geldt des te meer in ‘een sociale (en soms politieke) context waarin mogelijk ook andere kwesties spelen’ en gegeven de complexiteit van het te gebruiken normenkader (Zouridis en van der Vorm, 2013, 13).

De aanwezigheid van een opdrachtgever, de problematische context waarin het onderzoek wordt verricht, de problemen die zich kunnen voordoen in het onderzoeksproces en de positie van de onderzoeker ten opzichte van de opdrachtgever maken dat zulk integriteitsonderzoek verschilt van wetenschappelijk onderzoek.

Het is veelal de opdrachtgever die de onderzoeksvraag en de deelvragen formuleert, die de onderzoeker een financiële vergoeding geeft, die de termijn bepaalt waarbinnen het onderzoek moet zijn voltooid en die niet alleen een feitenrelaas verwacht, maar de onderzoeker ook om eenduidige conclusies vraagt. De context is vaak problematisch omdat er ernstige conflicten kunnen bestaan, de verdenkingen omstreden zijn, de onderzoeksvraag zelf al punt van discussie kan zijn en medewerking aan het onderzoek onzeker is.

Waar in wetenschappelijk onderzoek een onderzoeker zelf de onderzoeksvragen formuleert, is dat bij zulk onderzoek dus vaak niet het geval. Dit kan beperkingen opleveren in de verzameling, verwerking en analyse van informatie en het type conclusies dat kan worden getrokken. Dat leidt ertoe dat onderzoekers wantrouwend zouden moeten staan ten opzichte van hun opdrachtgever en zij consequenties zouden moeten trekken uit een gebrek aan medewerking of te beperkte vraagstelling (cf. Zouridis & van der Vorm, 2013).

BIOS heeft een lijst met uitgangspunten opgesteld om de kwaliteit van integriteitsonderzoek te waarborgen. Het gaat daarbij allereerst om de professionaliteit van de onderzoeker. De expertise en ervaring van de onderzoeker moet afgestemd zijn op de feitelijke onderzoeksvraag, de onderzoeker heeft kennis van en voeling met de bestuurlijk-politieke context, met voldoende aandacht voor de media en het politieke krachtenveld, en weet rekening te houden met de publieke context. Het onderzoek moet zonder willekeur worden verricht, subsidiair en proportioneel zijn - dat wil zeggen zo licht mogelijk - alsmede kostenefficiënt (cf. BIOS, 2014, pp. 16-17).

Integriteitsonderzoek wordt door onderzoekers van verschillende pluimage gedaan. Het kan bijvoorbeeld een organisatieadviesbureau zijn, een bestuurskundige, een accountant, of een organisatie waarin vooral mensen werkzaam zijn met een achtergrond als rechercheur. Voor elk van die onderzoekers gelden de criteria van hun eigen beroepsgroep. Als een accountant het integriteitsonderzoek uitvoert, geldt de gedragscode van de raad van accountants. Het basisprincipe van waaruit zij hun opdracht dienen vorm te geven is dat zij verantwoordelijk dienen te handelen in het algemeen belang. Hun verantwoordelijkheid is niet beperkt tot het behartigen van het belang van een individuele cliënt of werkgever. Als criteria staan professionaliteit, betrouwbaarheid, zorgvuldigheid, onafhankelijkheid, objectiviteit, integriteit en vertrouwelijkheid centraal. Deze criteria hebben binnen deze beroepsgroep een eigen betekenis. Zo betekent objectiviteit binnen deze beroepsgroep dat de accountant zich bij zijn afwegingen niet ongepast laat beïnvloeden en professionaliteit dat de accountant zich ‘onthoudt van elk handelen of nalaten waarvan hij weet of behoort te weten dat dit de beroepsgroep in diskrediet brengt of kan brengen’ (NIVRA, n.d.).

Het is mede daarom dat Zouridis en van der Vorm stellen dat ‘Eenduidige en gedeelde kwaliteitsmaatstaven en protocollen voor integriteitsonderzoek ontbreken’ en ‘Er zijn weliswaar gedragscodes en onderzoekreglementen, maar er zijn geen algemeen gedeelde standaarden en kwaliteitsmaatstaven voor integriteitsonderzoek. Dat maakt de beoordeling van concreet onderzoek voor opdrachtgevers lastig.’ (Zouridis en van der Vorm, 2013, pp. 43-44).

Hieronder wordt beschreven wat de gevolgen zijn van de specifieke kenmerken van integriteitsonderzoek voor de validiteit en betrouwbaarheid ervan, dat wil zeggen in hoeverre in de inhoud van zulk onderzoek indicaties zijn te vinden dat de kwaliteitseisen gesteld vanuit wetenschap en praktijk met elkaar botsen. Dat zou mag worden verwacht gegeven de nadruk die in elk protocol en handreiking is te vinden over de complexiteit van de context waarin zulk onderzoek wordt uitgevoerd en hoe belangrijk bijvoorbeeld de professionaliteit van de onderzoeker is. Door specifieke kenmerken van de context zou bijvoorbeeld de onderzoeksvraag zodanig geformuleerd kunnen zijn dat het tot conclusies leidt die niet of minder passend zijn voor waar het onderliggende probleem om vraagt. Ook kan de te gebruiken onderzoeksmethode, de wijze waarop de onderzoeker tracht de feiten boven tafel te krijgen, niet of minder adequaat zijn bij de onderliggende vraag. Daarbij is de validiteit in het geding, dat wil zeggen de vraag of wordt gemeten wat men beoogt te meten. De analyses, conclusies en verslaglegging kunnen gegeven de invloed van de context en de gevoeligheid van de onderzoeker daarvoor, tot uitkomsten leiden waardoor de betrouwbaarheid in het geding komt. Dat wil zeggen dat dit kan variëren afhankelijk van wie het onderzoek uitvoert. Een voorbeeld daarvan staat in box 1.

Box 1. *Een illustratie van in het geding zijnde validiteit en betrouwbaarheid*

De context: Een Barnevelds raadslid wordt verweten dat hij een kavel met huis erop had gekocht en zich daarna als raadslid met de besluitvorming over het bestemmingsplan dienaangaande had bemoeid. De vorige eigenaar had geen toestemming gekregen van de gemeente om de kavel te splitsen, terwijl dat na de verkoop opeens wel mogelijk leek door een wijziging in het bestemmingsplan. De eerdere eigenaar zag daar de hand van het betreffende raadslid in en stelde als verwijtende vraag in een zienswijze op de voorgenomen wijziging van het bestemmingsplan of het raadslid wellicht met voorkennis had gekocht en of het raadslid invloed had gehad op het voorstel tot wijziging. Het raadslid verzocht de indiener dringend om deze beschuldigingen in te trekken. Daarover klaagde de indiener en de burgemeester gaf daarop de opdracht tot het instellen van een integriteitsonderzoek naar de handelwijze van het raadslid.

De onderzoeksvraag: Deze formuleert de onderzoeker als volgt: “is er in de behandeling van de aanvraag van de ruimtelijke ontwikkeling rechtmatig en integer gehandeld? “ met daarbij twee deelvragen, namelijk of de gemeente bij de behandeling van de bouwvergunning met twee maten heeft gemeten? En of het raadslid integer heeft gehandeld (Van den Heuvel, 2015a, p. 4).

Conclusies van het onderzoek: De gemeente valt niets te verwijten. Wijzigingen in het bestemmingsplan zijn een normale zaak

gegeven bijvoorbeeld een veranderende context. Het raadslid wordt door de onderzoeker bekritiseerd over zijn reactie op de zienswijze van de indiener. Dat had volgens de onderzoeker niet gemogen, omdat daarmee het besluitvormingsproces was beïnvloed door een raadslid met een persoonlijk belang in de zaak.
Gevolg: Het raadslid wordt door zijn partij ontoelaatbaar gedrag verweten, de gemeenteraad neemt een motie van treurnis tegen hem aan en hij wordt onhoudbaar binnen zijn partij.
Een contra-expertise: Het raadslid laat een contra-expertise uitvoeren. Die komt op basis van hetzelfde feitenrelaas tot een geheel andere conclusie, namelijk dat het betreffende raadslid niets te verwijten is. Iedereen was gebaat bij de wijziging in het bestemmingsplan. Het raadslid had de indiener er enkel van proberen te overtuigen dat het beter is om de beschuldigingen aan zijn adres terug te nemen en eigenlijk was de burgemeester – de opdrachtgever van het eerste onderzoek – verantwoordelijk. Die had het raadslid moeten bijstaan en omdat dit achterwege bleef moest het raadslid wel zelf in actie komen (Elzinga & De Haan, 2015).

Het gaat er in dit voorbeeld niet om de vraag welke onderzoeker gelijk heeft. Voor het antwoord daarop is onderzoek naar het oorspronkelijke, maar uit de opdracht weggelaten, probleem nodig, namelijk of er indicaties waren dat het raadslid inderdaad voorkennis had en of hij invloed had uitgeoefend op het gemeentelijk apparaat bij de wijziging van het bestemmingsplan. Het voorbeeld dient om te illustreren hoe bij twee onderzoeken naar eenzelfde zaak de validiteit en betrouwbaarheid in het geding kan zijn. De validiteit omdat twijfelachtig is of is vastgesteld wat had moeten worden vastgesteld, en de betrouwbaarheid omdat klaarblijkelijk verschillende onderzoekers op basis van dezelfde feiten tot tegenovergestelde bevindingen komen. In dit geval is dat gerelateerd aan de context waarin beide onderzoeken werden uitgevoerd, de onderscheiden normenkaders en opdrachtgevers.

In het vervolg van dit artikel worden meer van dergelijke voorbeelden gepresenteerd. Dat wordt gestructureerd door de problemen per onderzoeksfase te onderscheiden. In veel handreikingen (zie bijvoorbeeld Protocol integriteitonderzoeken Gemeente Amsterdam (2006); BIOS, handreiking onderzoek integriteitsschendingen politieke ambtsdragers (2014)) worden voor integriteitsonderzoek minimaal de volgende drie fasen onderscheiden:
1) De fase voorafgaand aan het onderzoek, inclusief de aanleiding voor het onderzoek en de formulering van de opdracht en onderzoeksvraag,
2) De fase van de uitvoering van het feitenonderzoek, met name de wijze van dataverzameling, inclusief de informatieverstrekking door opdrachtgever en opdrachtnemer de medewerking aan het onderzoek door betrokkenen, en de bepaling van de te gebruiken onderzoeksmethode (interviews, documentanalyse, observatie, doorzoeken van de werkomgeving of telecommunicatie),
3) de afronding/verslaglegging, inclusief de analyse en conclusies, hoor- en wederhoor, en de rapportage.

Deze drie fasen worden hieronder ook onderscheiden. Alvorens in te gaan op de problematiek in elk van deze fasen wordt eerst kort ingegaan op de gebruikte onderzoeksrapporten.

3 De data

In totaal zijn er 47 rapporten onderzocht, waarvan twee op dezelfde zaak betrekking hadden, maar met verschillende verdachten, een ander feitenrelaas en andere conclusies. De rapporten zijn alle geschreven tussen 2010 en 2015 en variëren sterk in aantal bladzijden (tussen de 4 en 248 bladzijden) en opzet. Ondanks dat vrijwel alle rapporten geheim of vertrouwelijk waren zij allemaal te vinden op het internet. In tabel 1 is te zien om welke soort rapporten het ging om welk type schendingen.

In een derde van de gevallen betrof het onderzoek naar de gedragingen van een specifieke persoon, in drie gevallen was er geen duidelijke verdachte op het moment dat het onderzoek van start ging, in 26 gevallen ging het om een onderzoek naar de cultuur, in twee gevallen om een contra-expertise en in een geval om een preventief onderzoek. Voor wat betreft de aard van de zaak ging het in 19 gevallen om vermeende belangenverstrengeling, in acht gevallen om financiële delicten, in vijf gevallen om het ‘lekken van vertrouwelijke informatie’ in vijf gevallen om ongewenste omgangsvormen, in drie gevallen om misbruik van bevoegdheden, in eveneens drie gevallen om onjuist gebruik van bedrijfsmiddelen, en in 11 gevallen was het onduidelijk wat de integriteitschending inhield. In tabel 1 zijn de aard van de verdachte en de aard van de integriteitschending tegen elkaar afgezet.

Soort rapport						**Totaal**
	Onderzoek specifiek persoon	Onderzoek schending zonder verdachte	Onderzoek cultuur	Contra-expertise	Preventief onderzoek	
Belangenverstrengeling	10	0	6	2	1	19
Financieel	0	1	7	0	0	8
Lekken van informatie	2	2	1	0	0	5
Ongewenste omgangsvormen	2	0	3	0	0	5
Misbruik van bevoegdheden	2	0	1	0	0	3
Misbruik bedrijfsmiddelen	1	0	2	0	0	3
Geen duidelijke schending	0	0	11	0	0	11
Totaal	15	3	26	2	1	47

Tabel 1: *Het onderliggende probleem in het integriteitsonderzoek*

Het grootste deel van de rapporten (65,6%) speelde zich af op gemeentelijk niveau. In 12,8% van de gevallen ging het om wethouders, in 4,3% om burgemeesters, en in 17,0% om leden of fracties van de gemeenteraad. In 8,5% van de rapporten ging om het college van burgemeesters en wethouders en 19,1% om de gemeente in het algemeen of werd er niet nader gespecificeerd wie in de gemeente verantwoordelijk was. Daarnaast speelt 6,4% zich af op provinciaal niveau, in één geval op nationaal niveau (tweede kamer), 10,6% rondom Zelfstandig bestuursorganen (ZBO) of stichtingen, 6,4% rondom specifieke overheidsprojecten en 6,4% gaan over ambtenaren of medewerkers. Opdrachtgevers op gemeentelijk niveau waren met name burgemeesters (19,1%) en de gemeenteraad (27,7%).

In 48,9% van de gevallen is de opdrachtgever ook de opsteller van de opdrachtformulering. In 42,6% van de rapporten wordt de opsteller daarvan niet expliciet genoemd. Daarnaast heeft in een derde van de rapporten de opdrachtgever de onderzoeksvragen opgesteld. In de helft van de gevallen wordt niet vermeld wie de onderzoeksvragen heeft opgesteld. In 53,2% van de rapporten werden knelpunten genoemd. De belangrijkste knelpunten zitten met name rondom de opdracht (31,9% van de rapporten), (aangeleverde) bronnen (29,8%) en medewerking (23,4%).

Om de rapporten te kunnen analyseren zijn zij gecodeerd aan de hand van een codeboek dat is opgesteld met als rode draad, de 'Handreiking: Onderzoek integriteitschendingen politiek ambtsdragers' van (BIOS, 2014). In die handreiking staat beschreven waaruit een integriteitsrapport hoort te bestaan. Daarna is dit aangevuld met bevindingen van (Zouridis & Vorm, 2013). Uiteindelijk bestond het codeboek uit 37 vragen. Hieronder worden de begrippen die het belangrijkst zijn voor de analyse geoperationaliseerd.

Ten eerste is non-respons gemeten. Non respons is het niet meewerken of niet reageren van betrokkenen. In dit onderzoek is met name gekeken hoe de onderzoekers de non-respons hebben geïnterpreteerd. Dit is gedaan aan de hand van vier categorieën. 1) Neutrale interpretatie, hierbij melden de onderzoekers de non-respons, maar wordt er geen informatie vermeld die de lezer kan beïnvloeden. 2) Negatieve interpretatie, hierbij melden de onderzoekers dat de non-respons een negatieve invloed heeft op het onderzoek. 3) Positieve interpretatie, de onderzoekers hebben er begrip voor dat iemand niet wil meewerken aan het onderzoek. En 4) Geen non-respons, de onderzoekers melden dat iedereen heeft meegewerkt.

Ten tweede is het feitenrelaas inclusief eventuele onzekerheden en onjuistheden gecodeerd. In vrijwel alle (95,7%) rapporten wordt een feitenrelaas gegeven. De meeste feitenreconstructies zijn gebaseerd op documentanalyses (91,1%) en/of interviews (88,9%). Om de onzekerheden en onjuistheden in het feitenrelaas te kunnen meten is de nadruk gelegd op deze twee elementen. Zaken zijn als onduidelijkheden of onzekerheden gerekend als bronnen niet konden worden terug gevonden, niet bestonden of de geïnterviewde betrokkenen elkaar tegen spraken (en hun gelijk niet kon worden bevestigd door andere bronnen). In dit onderzoek waren er 46,8% van de gevallen onduidelijkheden en/of onzekerheden te vinden in het feitenrelaas.

Ten derde is er gekeken of er hoor en wederhoor heeft plaatsgevonden. Dat betekent dat de betrokkenen en/of opdrachtgever konden reageren op het conceptrapport. De manier waarop hoor en wederhoor wordt ingericht verschilt (Zouridis & Van der Vorm, 2013, p. 39). Van alle rapporten die hier zijn geanalyseerd heeft slechts in iets meer dan de helft (55,3%) van de gevallen hoor en wederhoor plaatsgevonden. Daarnaast is er gekeken of uit het rapport duidelijk wordt wat er met de hoor en wederhoor is gedaan door de onderzoekers. Hiervoor zijn vijf categorieën gemaakt. In de eerste categorie wordt in het rapport duidelijk dat alle conclusies die het betrof zijn aangepast aan de hand van de feedback (2 rapporten). In de tweede categorie is een significant deel aangepast. Dit houdt in dat er na de feedback punten in het rapport zijn aangepast die cruciaal zijn voor het beantwoorden van de onderzoeksvragen (5 rapporten). In de derde categorie is slechts een gering deel in het rapport aangepast. Het gaat dan om zaken die niet relevant zijn voor de onderzoeksvragen (2 rapporten). De vierde categorie gaat om rapporten waarin niets is aangepast na het wederhoor (1 rapport) en in de vijfde categorie wordt niet duidelijk wat er is aangepast (15 rapporten).

Ten vierde is er gekeken of er drogredeneringen in de rapporten stonden. Drogredeneringen zijn argumenten die op het eerste gezicht kloppend lijken, maar zijn gebaseerd op verkeerde premissen. In dit onderzoek stonden in 12,8% van de rapporten zulke drogredenen.

Ten vijfde is in het onderzoek gekeken of data en moraliserende opmerkingen van elkaar gescheiden zijn. Dit is het geval wanneer onderzoekers wezen op wat in hun ogen juist was, zonder daar bij te vermelden dat het hun mening was en waarbij die mening als feit werd gepresenteerd. In iets minder dan een kwart (23,4%) van de rapporten werden data en moraliserende opmerkingen niet van elkaar gescheiden.

Als laatste is er gekeken of de conclusies houdbaar zijn. Hierbij is op vier zaken gelet. Ten eerste is er gekeken of de conclusies gebaseerd zijn op een adequate vergelijking van het feitenrelaas met het normenkader. Daarbij is er gekeken of er geen onzekerheden of onduidelijkheden in het feitenrelaas stonden, het feitenrelaas getoetst is aan een normenkader, of er geen drogredenen zijn gebruikt en of data en moraliserende waarden van elkaar gescheiden waren. Voldeed het rapport aan al deze zaken dan werd de conclusie houdbaar genoemd. In totaal kwam op basis van deze criteria slechts iets meer dan één op de vier integriteitsrapporten (27,7%) met een conclusie die op basis van deze criteria als houdbaar kan worden beschouwd.

Omdat dit percentage opvallend laag is, wordt hieronder een inhoudelijke analyse per onderzoeksfase gegeven. Daarin zal worden beschreven wat men zoal tegenkomt in elke fase en waar onderzoekers anders te werk gaan dan men zou verwachten. Dit wordt gevolgd door een analyse waarin de kenmerken van die onderzoekingen die wel tot een te rechtvaardigen conclusie kwamen worden vergeleken met de onderzoekingen waarin twijfels bestaan over de gerechtvaardigdheid van de conclusies.

4 Wat opvalt in elk van de onderzoeksfasen

4.1. Fase 1. Voorafgaand aan het onderzoek: het belang van de onderzoeksvraag

In vrijwel al het geanalyseerde onderzoek over integriteitsproblemen in het openbaar bestuur valt op dat de onderzoekers er in refereren aan de opdracht die zij hebben meegekregen en aan de beperkingen die dit met zich mee heeft gebracht. Daar is op zich niets mis mee, behalve als die opdrachtformulering reeds in belangrijke mate bepalend is voor de uitkomsten van het onderzoek. Dat was het geval in het onderzoek naar de integriteitsproblemen binnen de Provincie Noord Holland waarin de onderzoekers onder de noemer "Operatie schoon schip" de opdracht meekregen onderzoek te doen naar die incidenten waarbij bestuurders waren betrokken. Als de commissie daarna een oproep doet om zaken te melden, krijgt ze 122 meldingen waarvan vrijwel gelijk 46 afvallen omdat ze buiten deze opdracht vallen (bijv. over provinciale ambtenaren). De overige meldingen gaan voor het overgrote deel over één bestuurder (Hooijmaijers) die toch al was opgestapt. Een aantal andere meldingen is volgens de commissie te vaag, of onder de rechter. En als dan uit een overgebleven casus toch blijkt dat de toenmalige commissaris van de koningin zich in Zuid Afrika heeft laten fêteren door een Noord Hollandse vastgoedondernemer, wordt dat zonder verder commentaar in het rapport vermeld. Een ander voorbeeld betreft het rapport naar het declaratiegedrag van het college van Gedeputeerde Staten van de provincie Gelderland opgesteld door Deloitte. Het gaat daarbij om de creditcard uitgaven en ingediende declaraties. Echter wordt door de onderzoekers gesteld dat een accountantscontrole geen onderdeel uitmaakt van de onderzoeksopdracht en dat er daarom 'aan onze rapportage geen zekerheid kan worden ontleend omtrent de getrouwheid van de onderzochte (financiële) informatie' (Deloitte, n.d., pp. 8–9).

Het tweede dat opvalt in veel onderzoek naar de integriteit van bestuurders is de politieke bemoeienis ermee. Dit gebeurt bijvoorbeeld door een (oud-) politicus voorzitter van de onderzoekscommissie te maken of door het onderzoek te laten verrichten door personen met partijpolitieke banden. Politieke bemoeienis kwam ook tot uiting in het onderzoek naar Operatie Schoon Schip waarin de commissie meent naast de *bad practices*, ook *best practices* te moeten constateren in de provincie Noord Holland, omdat zo zegt de commissie zij beseft dat haar onderzoek en de bevindingen negatief kunnen afstralen op de VVD, omdat de verdachte Hooijmaijers lid van de VVD is (Van den Heuvel, Pheijffer, & De Vries, 2012, pp. 14–16;132).

Het ging volgens de commissie om het contrast tussen de declaraties van de ene en andere VVD gedeputeerde; het feit dat de ene VVD-er de andere in de Staten vaak van repliek diende; en het feit dat de ene VVD-er zich veel minder ontvankelijk toonde voor de vastgoedsector en projectontwikkelaars dan de andere. Letterlijk geciteerd: "Het betreft voorbeelden die de Commissie in haar onderzoek ziet als *best practices*." Het gaat wat ver om dit best practices te noemen en de enige reden waarom dit soort zaken in het onderzoek worden opgenomen lijkt politieke gevoeligheid i.c. politieke bemoeienis te zijn.

Het derde probleem is dat het steeds meer usance wordt in opdrachtonderzoek dat de opdrachtgever al inzage krijgt in het rapport voordat het wordt gepubliceerd en zelfs de mogelijkheid krijgt de uitkomsten van het onderzoek nog aan te passen. Enerzijds lijkt dat handig om eventuele foutjes tijdig uit het rapport te halen, maar anderzijds levert dit het risico op dat de opdrachtgever de conclusies gaat aanpassen of de onderzoekers vraagt wat meer aandacht aan de positieve kanten en minder aan de negatieve kanten van de zaak te besteden. Zo zag de burgemeester van Roermond er geen bewaar in de zaak rond Van Rey de onderzoekers te vragen om het woord "Belangenverstrengeling" uit het rapport te verwijderen. De NRC maakte daarop duidelijk dat dit waarschijnlijk niet het enige verzoek is waar de onderzoekers mee te maken kregen (Dohmen, 2012).

In een zaak in het Westland gaat het om een fractie in de raad die kritiek uit op de handelwijze van de gemeentelijke accountant die zowel adviezen aan de gemeente geeft als de jaarrekening controleert. De kernvraag die de onderzoeker daaruit eerst afleidt is of het juist en opportuun was dat de accountant van 'fout optreden' werd beschuldigd en of die beschuldiging, gezien de inrichting en het functioneren van de democratie, legitiem was (Van den Heuvel, 2015b, p. 5). Omdat de fractie echter een klacht heeft neergelegd bij de accountantskamer, staat even later in het onderzoek, 'Nu de zaak onder de rechter is, mag de onderzoeker in dit rapport voor de gemeente niet op de stoel van de rechter gaan zitten.

Hij kan zich geen uitspraken over de accountant veroorloven.' Afgezien van de vraag of dit een adequaat argument is, wordt de overblijvende vraag volgens de onderzoeker, of 'het optreden van de fractie met haar klacht tegen de accountant, mede in het licht van het acteren van de raad, het presidium, de voorzitter van de gemeenteraad en de griffier, in deze zaak correct is geweest, gezien de wettelijke en daarvan afgeleide regels.' Om even later echter met weer tot de eerste onderzoeksvraag te komen, namelijk 'of de gemeente de accountant strak genoeg aan de teugels had om controle te hebben en te houden op adviesopdrachten en ontvangen declaraties.' Waarna uiteindelijk toch de onderzoeksopdracht wordt, 'te onderzoeken: de handelwijze van de fractie rond de problemen met de accountant'. Hierna wordt de uiteindelijke opdrachtformulering de handelwijze van de fractie, waar het gaat om: a. het schenden van de vertrouwelijkheid van een besloten presidiumvergadering 'door elementen daaruit voor publiciteitsdoeleinden te gebruiken om de vooraankondiging van de klacht te legitimeren', b. het passeren van de raad 'als hoogste bestuursorgaan in ons democratisch bestel, als hoeder van het algemeen belang en als opdrachtgever van de accountant' (idem, p.12). Soms zit het venijn van een onderzoeksvraag in kleine woordjes. De eerste vraag werd in het onderzoek namelijk veranderd van 'schenden van vertrouwelijkheid' naar 'het bruuskeren van de vertrouwelijkheid'. Het betreffende raadslid had aanwezig willen zijn bij een besloten vergadering tussen de accountant en het presidium. Daarop werd die vergadering afgeblazen. Van schenden van vertrouwelijkheid kon geen sprake zijn.

Te zien is een geleidelijke verschuiving van de onderzoeksvraag van een vraag waarin het handelen van de accountant centraal staat naar een onderzoeksvraag waarin alleen het kwalijk bevonden handelen van de fractie die de accountant had aangeklaagd wordt onderzocht.

Nu duidelijk is dat integriteitsonderzoek zelf problematisch kan zijn, wordt in het volgende onderzocht of dat een incidenteel of een structureel probleem is. Daartoe wordt eerst aangegeven hoe in de rapporten informatie is verzameld en hoe de inhoud daarvan is geanalyseerd. Dat wordt in de hiernavolgende paragraaf gedaan.

4.2. De dataverzameling

In het overgrote deel van de integriteitsonderzoeken baseren de onderzoekers zich op een combinatie van interviews en document analyses. In tabel 2 wordt uiteengezet welke methoden sporadisch ook worden gebruikt.

Gebruikte onderzoeksmethode	Aantal	Percentage
Document analyse	42	89,4%
Interviews	41	87,2%
Onderzoek (tele-) communicatie	12	25,5%
Notitie / advies expert	7	14,9%
Operationalisatie/ literatuurstudie	7	14,9%
Schriftelijke vragen/ reactie	7	14,9%
Raadpleging eerdere onderzoeken	6	12,8%
Raadpleging openbare bronnen	5	10,6%
Enquête	4	8,5%
Opstellen tijdlijn	3	6,4%
Informatie bewaring en verzameling	3	6,4%
Netwerkanalyse	2	4,3%
Onderzoek van werkplek	1	2,1%
Statistisch onderzoek / Data analyse	1	2,1%
Observatie	0	0,0%
Wordt niet genoemd	3	6,4%

Tabel 2: *Onderzoeksmethoden van de rapporten*

De twee meest gebruikte onderzoeksmethoden kunnen een aantal problemen met zich meebrengen. Een veel voorkomend probleem met interviews is dat betrokkenen zich gebeurtenissen anders herinneren en soms tegengestelde herinneringen hebben. Het is dan vaak lastig voor de onderzoekers om na te gaan wie er gelijk heeft. Het is in dat geval het ene woord tegen het anderen. Een voorbeeld hiervan is te vinden in het rapport van BING over de deelgemeente Fijenoord (2013). Hierin wordt gekeken of er mogelijke belangenverstrengelingen zijn van deelraadsleden bij stichtingen en projecten. Een aantal geïnterviewde geeft aan dat er in het verleden sprake is geweest van belangenverstrengeling, maar door de verdachten hiervan wordt dat in alle toonaarden ontkend. Door de onderzoekscommissie wordt niet aangegeven of en dat er bronnen zijn om te ondersteunen wie er gelijk heeft.

Dit lijkt toch een issue te moeten zijn voor het feitenrelaas en de conclusies van het rapport. Dit wordt echter door de onderzoekers in het geheel niet aangemerkt als knelpunt. Dit blijkt structureel. Er wordt maar in twee rapporten vermeld dat het ene woord tegen het andere een knelpunt was voor het trekken van conclusies, terwijl het in veel meer rapporten een zichtbaar probleem is.

Daarnaast is in de rapporten te zien dat mensen van wie de onderzoekers graag informatie hadden willen hebben soms niet meewerken aan het onderzoek. Dat is in elk onderzoek een te verwachten probleem, maar in integriteitsonderzoek spelen enkele daarvoor specifieke zaken. Soms weigeren

betrokkenen expliciet om aan het onderzoek mee te werken, omdat ze het oneens zijn met de opdrachtformulering, of omdat ze geen vertrouwen hebben in de onafhankelijkheid van het onderzoek. Voor zover het gaat om enquêtes onder het personeel is er de gebruikelijke non-respons die soms hoger en soms lager uitvalt.

Opvallend in integriteitsonderzoek is hoe de onderzoekers hiermee omgaan:

De eerste mogelijkheid is om het niet-meewerken te laten voor wat het is, en de mensen die wel deelnemen toch als een representatieve steekproef van de gehele populatie te zien. Dat is onproblematisch als de mensen die wel deelnemen aan het onderzoek inderdaad representatief zijn voor de gehele. Als echter de non-respondenten een specifieke groep vormen, waarvan verwacht mag worden dat hun opvattingen afwijken van de respondenten, dan vormt generalisatie over het geheel wel een probleem. Dat was bijvoorbeeld te zien in het onderzoek van BING naar de seks-rel in Wassenaar (2012). Allen die de zaak hadden aangekaart weigerden vanwege hun probleem met de formulering van de onderzoeksvraag mee te werken aan het onderzoek. De enigen die meewerkten aan het onderzoek waren degenen die ontkenden dat er sprake was geweest van onbehoorlijk gedrag. In een dergelijk geval kunnen de opvattingen van degenen die wel hebben deelgenomen, uiteraard niet zonder meer als representatief voor het geheel worden gezien. Bij de presentatie van het rapport werd echter door de onderzoekers geconcludeerd dat niemand had bevestigd dat er sprake was geweest van onbehoorlijk handelen en dat het plaatsvinden ervan daarom zeer onwaarschijnlijk zou zijn.

Een tweede discutabele interpretatie van niet meewerken, die is te zien in onderzoeksrapporten, is om degenen die niet meewerken aan het onderzoek gelijk als verdacht te beschouwen. Dat is te zien in een onderzoek van bureau Hoffmann naar een lek uit een vertrouwelijk overleg in de gemeenteraad van Urk (2015). Een raadslid dat vanuit privacyoverwegingen de onderzoekers inzage in zijn e-mails weigerde werd in het uiteindelijke rapport bestempeld met de frase ‘merkwaardig gedrag’, waarmee tijdens de bespreking van het rapport in de raad gelijk de beschuldigende vinger naar dat raadslid ging.

Een derde interpretatie van non-respons is deze te interpreteren als ondersteunend aan de conclusie te zien. In onderzoek naar het eventuele bestaan van een angstcultuur waarvoor een directeur of gemeentesecretaris verantwoordelijk is, wordt de non-respons dan bijvoorbeeld geïnterpreteerd als indicatief voor het bestaan van die angstcultuur waardoor mensen niet voor hun mening uit zouden durven komen.

Naast de problemen met interviews brengt ook het doen van een document analyse problemen met zich mee. Veel rapporten baseren zich op informatie die aangeleverd wordt door derden zonder deze zelfstandig te verifiëren. In een aantal rapporten waarin declaraties centraal staan wordt aangegeven dat er geen accountantscontrole is uitgevoerd en er rondom de financiële aspecten van de conclusie dus geen zekerheid kan worden gegeven. Ook in het rapport van KPMG over de relatie tussen het Baggerfestival en de gemeente Sliedrecht wordt gesteld dat ze de rapportage hebben op gesteld aan de hand van schriftelijk en mondelingen informatie die aan hun verstrekt is. Hierover zegt KPMG ‘Dit impliceert dat de juistheid en volledigheid van de in dit rapport opgenomen informatie afhankelijk is deze ter beschikking gestelde mondelinge en schriftelijke informatie. Wij hebben niet zelfstandig de juistheid en volledigheid van de verstrekte informatie onderzocht, anders dan beschreven in dit rapport. De aan ons verstrekte mededelingen door derden zijn verstrekt op basis van vrijwilligheid en in ons onderzoek betrokken met als uitgangspunt dat deze niet onjuist of misleidend zijn’ (2011, p. 3). De basis van vrijwilligheid neemt echter niet weg dat de betrokkenen in dit onderzoek hun eigen belangen kunnen hebben.

4.3. De analyse en redenering naar een conclusie

Voor de analyse zijn het feitenrelaas en het normenkader het belangrijkst. In de analyse worden de feiten namelijk normenkader, dat vaak is opgesteld aan de hand van gedragscodes en regelgeving, om te zien of er daadwerkelijk sprake was een integriteitsschending.

Voor verreweg de meeste rapporten (80,9%) is er een normenkader opgesteld waaraan de feiten worden getoetst. Wat hierbij opvalt is het grote aantal bronnen dat gebruikt is voor deze normenkaders. In totaal is er in de 38 rapporten met normenkaders gebruik gemaakt van maar liefst 74 verschillende soorten bronnen. De keuze voor bepaalde bronnen, maar ook de manier waarop de bronnen in het normenkader

geïnterpreteerd worden, kan de analyse en conclusie beïnvloeden. Gedragscodes en andere wet- en regelgeving zijn namelijk vaak breed toepasbaar en de manier waarop onderzoekers deze interpreteren kan dan ook verschillen. Het duidelijkste voorbeeld hiervan is omschreven in box 1.

In vrijwel alle (95,7%) rapporten is een feitenrelaas opgesteld. Wanneer dit is gebaseerd op incomplete of verkeerde informatie ontstaan er onzekerheden en/of onjuistheden in dit feitenrelaas. Redenen hiervoor zijn deels omschreven in de vorige paragraaf. Onzekerheden en onjuistheden in het feitenrelaas hebben uiteraard gevolgen voor de conclusie. De conclusie dient namelijk aan te sluiten op een vergelijking van het feitenrelaas met het normeringskader. Wanneer in het feitenrelaas onjuiste of incomplete informatie zichtbaar is, moet dat gevolgen hebben voor de houdbaarheid van de conclusie.

Na de opstelling van het concept rapport wordt door de onderzoekers een mogelijkheid aangeboden tot reactie, hoor en wederhoor, maar uit dit onderzoek blijkt dat dit slechts in iets meer dan de helft van de rapporten (55,3%) het geval is. En ook al heeft hoor en wederhoor plaatsgevonden, dan is nog in 57,7% van die rapporten niet duidelijk wat er met de reactie van de wederhoor is gebeurd. Dat betekent dat er slechts 10 rapporten in dit onderzoek zijn gevonden waarin duidelijk is dat er hoor en wederhoor heeft plaatsgevonden en dat op basis daarvan iets is aangepast. In enkele gevallen is een transcriptie van het gehele hoor en wederhoor opgenomen in het onderzoek. Daaruit blijkt dat in het grootste deel daarvan het wederhoor de bevindingen van de onderzoekers tegenspreekt.

5. Een vergelijking van de rapporten met gerechtvaardigde en niet-gerechtvaardigde conclusies

Het is teleurstellend dat van de 47 rapporten slechts 27,7% tot een houdbare conclusie leidt gegeven de hierboven genoemde criteria. Vier rapporten hebben geen conclusie en worden niet meegenomen in deze analyse. Eerder in dit onderzoek is al vermeld dat om tot een houdbare conclusie te komen er aan vier voorwaarden moet worden voldaan (normenkader, geen onduidelijkheden of onzekerheden in het feitenrelaas, geen drogredenen, data en moraliserende opmerkingen zijn gescheiden). Tabel 3 omschrijft hoeveel van deze zaken mis gingen in de rapporten en op welk van deze punten het mis ging.

		Op welke punten ging het mis				Totaal
		Normenkader	Onzekerheden en/of onduidelijkheden in feitenrelaas	Drogredenen	Data en moraliserende waarden niet gescheiden	
Op hoeveel punten ging het mis	Eén	4	16	0	3	23
	Twee	3	4	2	3	6
	Drie	1	4	3	4	4
	Vier	1	1	1	1	1
Totaal		9	25	6	11	34

Tabel 3: *Waar ging het mis in de rapporten met onhoudbare conclusies*

In de meeste gevallen gaat het mis op één punt en dan met name in het feitenrelaas. Wat daarnaast opvallend is, is dat in vijf van de zes rapporten waarin drogredenen worden gebruikt ook sprake is van data en moraliserende waarden die niet van elkaar gescheiden zijn.

Wanneer de rapporten met houdbare en niet houdbare conclusies met elkaar worden vergeleken vallen een aantal zaken op. Ten eerste is te zien dat hoe dikker het rapport is, hoe vaker de conclusies discutabel zijn. Er zijn geen rapporten met houdbare conclusies met meer dan 150 pagina's. De dikke rapporten hebben twijfelachtige conclusies aangezien er vaak onduidelijkheden en onzekerheden in het feitenrelaas staan. In het dikste rapport (248 pagina's) zijn daarnaast ook nog drogredenen te vinden en zijn data en moraliserende waarden niet van elkaar gescheiden. De meeste rapporten bestaan uit 4 tot 60 pagina's en in die categorie zijn ook de meeste rapporten te vinden met houdbare conclusies.

Het tweede wat opvalt wanneer de rapporten met elkaar worden vergeleken is dat er in de rapporten met houdbare conclusies veel minder sprake is van non respons. In 6 van de 13 rapporten met

houdbare conclusies wordt aangegeven dat iedereen aan het onderzoek heeft meegewerkt, dit in tegenstelling tot 2 van de 30 rapporten met discutabele conclusies. Dit kan deels worden verklaard doordat in de rapporten waarin er geen sprake was van non respons ook minder vaak onzekerheden en onduidelijkheden in het feitenrelaas stonden. In slechts één van de acht rapporten zonder non respons waren er onduidelijkheden en/of onzekerheden in het feitenrelaas te vinden. In het overgrote deel van de rapporten met onhoudbare conclusies (19 van de 30) wordt niets vermeld over non respons.

Een derde punt dat opvalt is te zien in het normenkader. In slechts negen rapporten is er geen sprake van een normenkader. Dat betekent dat er in 38 rapporten wel een normenkader te vinden is. Van die 38 rapporten die wel expliciet uitgaan van een normenkader komen er toch nog 22 tot een discutabele conclusie. Een groot verschil is te zien tussen rapporten waarin het normenkader is gebaseerd op wetenschappelijke literatuur en operationalisaties – waarvan meer dan de helft tot een houdbare conclusie leidt- en rapporten met discutabele conclusies waarbij slechts vier van die 22 rapporten het normenkader op wetenschappelijke literatuur heeft gebaseerd.

Ten vierde valt op dat wanneer het rapport is opgesteld (met behulp van) een interne opdrachtnemer het veel vaker een discutabele conclusie heeft. Slechts één van de tien rapporten waarin het onderzoek is uitgevoerd door een interne instantie leidt tot een houdbare conclusie. Echter is in tabel 4 te zien dat veel andere opdrachtnemers het niet veel beter deden. Rapporten uitgevoerd door hoogleraren en bestuurs- en organisatieadviesbureaus hadden het vaakst een houdbare conclusie.

		Houdbare conclusie			Totaal
		Niet van toepassing (geen conclusie)	Ja	Nee	
Opdrachtnemer	Universiteit/ Hoogleraar	0	7	6	13
	(Samenwerking met) interne opdrachtnemer	1	1	8	10
	Bestuurs- en organisatieadviesbureaus	0	3	4	7
	Instanties voor integriteitsonderzoek en integriteitsbeleid	1	1	4	6
	(Forensische) accountants	1	1	3	5
	Juridisch dienstverleners en advocaten	1	0	3	4
	Bedrijfs- en particuliere recherchebureaus	0	0	2	2
Totaal		4	13	30	47

Tabel 4: *Welke opdrachtnemers hadden houdbare conclusies in hun rapporten*

6. De uitkomsten van het wederhoor met de onderzoekers

Een concept van dit artikel is verstuurd aan de onderzoekers van de integriteitsrapporten en zij zijn in de gelegenheid gesteld om hierop te reageren. Hierop kwamen uiteenlopende reacties. Aan de hand van de wederhoor zijn een aantal punten in dit artikel aangepast. Hieronder staan de belangrijkste punten uit de reacties op dit artikel.

In verschillende reacties werd aangegeven dat het onderwerp van het artikel belangrijk en interessant was.

BMC Onderzoek merkt op dat er in de praktijk vaak te snel een integriteitsonderzoek wordt ingesteld. Ook al blijkt er uiteindelijk geen sprake te zijn van een schending, dan nog is de naam van de bestuurder beschadigd. Ze onderschrijven de noodzaak van zorgvuldigheid. Volgens BMC zijn er niet alleen standaarden nodig voor onderzoek, maar ook standaarden voor wanneer er onderzoek zou moeten worden gedaan.

De standaarden voor onderzoek en verscheidenheid in het onderzoek zijn komen terug in de reactie van Capra Advocaten (Capra) en IRS Forensic Investigation & Integrity (IRS). Zij achten het bijzonder dat in de uitkomsten van dit onderzoek blijkt dat hoogleraren en organisatie- en adviesbureaus de meeste houdbare conclusies hebben. Capra vindt dit opvallend aangezien zij in de praktijk zien dat bestuurskundige onderzoekers een andere invalshoek hebben dan juristen. Daarnaast stelt Capra dat het een probleem is dat onderzoekers niet altijd bekend zijn met de relevante juristprudentie.

IRS vindt het verassend dat hoogleraren en organisatie- en adviesbureaus vaker houdbare conclusies hebben, aangezien deze instanties vaak geen vergunning hebben voor dit soort onderzoek. Deze vergunning is verplicht volgens de Wet op de Particuliere Beveiligingsorganisaties en Recherchebururo' s. Volgens IRS is dit een economisch delict en het verbaast hen dan ook dat dit niet in het artikel is genoemd. De auteur van dit artikel merkt echter op dat in de handreiking van BIOS in een voetnoot het volgende wordt vermeld, "Er bestaan voor het uitvoeren van integriteitsonderzoek nog geen kwaliteitskeurmerken, certificaten/keurmerken of branchenormen" (2014, p. 19).

De verscheidenheid in de onderzoeken wordt ook aangekaart in de reactie van professor dr. Huberts. Hij vraagt zich af of alle integriteitsonderzoeken wel aan de criteria, die in dit artikel worden genoemd, moeten voldoen of dat dit kan verschillen per type integriteitsonderzoek. Daarnaast stelt hij dat de criteria die in dit artikel worden gehanteerd niet op alle punten even duidelijk zijn ingevuld. Dit punt wordt ook genoemd in de reactie van Capra. Ook wordt er zowel door professor dr. Huberts als Berenschot kritiek geuit op de toon van het artikel. Met name de laatste alinea van de conclusie 'Barbertje moet hangen' wordt als een te forse uitspraak beschouwd.

Professor mr. Tak merkt in zijn reactie op dat voordat er word bepaald of integriteit is geschonden, het van belang is om te vast te stellen wat integriteit is. Wanneer dit niet wordt gedaan is er het risico dat de intrinsieke aard van integriteit wordt miskend en men blijft steken in de vorm en techniek. Hij bekritiseert ook de aanpak zoals in dit artikel gepresenteerd als een voorbeeld van de huidige tijdgeest waarin wordt gestreefd naar standaardisering en objectivering.

Andere reacties van onderzoekers gingen meer in op het onderzoek dat zij zelf doen. Zo stellen Hordijk & Hordijk dat zij nooit een conclusie formuleren en nooit het woord integriteit gebruiken.

Professor dr. Van den Heuvel wijst op het feit dat hij in de rechtszaak, die tegen zijn rapport werd aangespannen door het Barnevelds raadslid (box 1), als onderzoeker in het gelijk is gesteld.

Hoffmann Bedrijfsrecherche wijst er in hun reactie op dit artikel op dat er in hun onderzoek naar de gemeente Urk (2015) nooit een schuldige voor het lek is aangewezen en dat dit aan het begin van hun onderzoek al wordt vermeld. Zij herkennen zich er niet in dat het niet meewerken als verdacht wordt beschouwd, hoewel zij dat in hun onderzoek benoemen als 'merkwaardig gedrag', maar zagen dit wel als belemmerend voor het onderzoek.

KAFI Integrity geeft aan in het onderzoek 'Uit Balans' (2014) geen normenkader te gebruiken aangezien het een feitenonderzoek betrof waarbij per onderzoeksvraag werd opgesteld of het feit nader onderzocht moest worden. Daarnaast geven zij aan dat het in de praktijk gebruikelijk is dat de opdrachtgever de onderzoeksvragen formuleert en dat het onderzoeksbureau ze waar nodig bijstelt en weer voorlegt aan de opdrachtgever.

Van de onderzoekers die aan een universiteit afkomstig waren, vertelde één dat hij nauwkeurig aan de opdracht had en niet meer kon worden verwacht. Een ander wilde graag de specifieke analyses over elk rapport en vooral het eigen rapport zien en vond dat de nadruk nogal eenzijdig lag op de kritiek op zulke rapporten.

Een onderzoeksbureau dat veel van dit soort rapporten uitbrengt was gezien haar reactie geheel 'not amused'. Zij wilde niet dat haar reactie met naam en toenaam zou worden genoemd in deze paragraaf

met weerwoord en had enkel grote twijfels, hoewel niet echt onderbouwd, bij de kwaliteit van dit onderzoek.

Dit alles geeft aan dat er belangrijke verschillen van inzicht bestaan tussen de onderzoekers die zulke integriteitsrapporten maken en overtuigd zijn van hun zorgvuldigheid en de schrijver van dit artikel die daarin belangrijke onzorgvuldigheden ziet. Wel blijkt dat veel van de onderzoekers zelf ook de problemen in het onderzoek wel zien, maar dan vooral in het onderzoek van anderen; dat zij het doorslaggevende belang van de opdracht die zij krijgen inzien; en dat ook zij voorstander zijn van een inhoudelijk normenkader, naast het reeds bestaande procedurele onderzoekskader voor zulk onderzoek. Daarbij variëren zij er in of dat inhoudelijke normenkader een gestandaardiseerd normenkader moet zijn,. In de navolgende conclusies zal worden aangegeven dat dit ook volgens de schrijver gewenst is.

7. Conclusies

In dit artikel is onderzoek gedaan naar de kwaliteit van integriteitsrapporten. Het doen van integriteitsonderzoek kan lastig zijn aangezien integriteit een gevoelig onderwerp is en het heeft vergaande gevolgen heeft voor de persoon waarnaar onderzoek wordt gedaan. Het is dan ook van belang dat hier zorgvuldig mee wordt omgegaan. Wanneer er echter naar integriteitsonderzoeken wordt gekeken kan een sterke verscheidenheid worden gezien in kwaliteit van de onderzoeken. Dit komt mede doordat er geen standaard is voor integriteitsonderzoeken en de onderzoekers verschillende achtergronden hebben. In dit artikel stond daarom de volgende vraag centraal: In hoeverre voldoet de inhoud van integriteitsonderzoek in al haar fasen aan de kwaliteitseisen van de wetenschap?

Om deze vraag te kunnen beantwoorden zijn 47 integriteitsrapporten geanalyseerd. Deze rapporten zijn op verschillende punten getoetst aan de hand van een codeboek dat is opgesteld met behulp van de handreiking van BIOS en een artikel Zouridis en Van der Vorm daarover.

In het onderzoek is te zien dat er in alle drie de fasen van het integriteitsonderzoeken fouten worden gemaakt. Het begint al bij de opstelling van de opdracht en de bemoeienis van de opdrachtgever. Tijdens het doen van het onderzoek liggen de problemen met name in de medewerking van de betrokkenen. Onderzoeken zijn grotendeels gebaseerd op document analyse en interviews en daarmee afhankelijk van de kwaliteit van de respons in de interviews. Tegenstellingen in de respons en non respons leiden tot onjuistheden en onzekerheden in het feitenrelaas. In de derde fase wordt vaak niet vermeld of er wederhoor is gepleegd en welke uitkomsten deze had. Daarnaast zijn de normenkaders divers en vaak breed te interpreteren waardoor de conclusies van de onderzoekers kunnen verschillen. In een aantal artikelen waren tevens drogredenen en moraliserende opmerkingen te vinden.

Wanneer een rapport niet aan de vier voorwaarden voldeed kon worden gesteld werd de conclusies er van twijfelachtig of niet houdbaar zijn. Deze vier voorwaarden betroffen onjuistheden en onzekerheden in het feitenrelaas, geen toetsing aan de hand van een normenkader, het voorkomen van drogredenen, en het niet scheiden van data en moraliserende opmerkingen. Het toepassen van die criteria leidde er uiteindelijk toe dat slechts van 27.7% van de rapporten kon worden gesteld dat de conclusie er in houdbaar is.

Gegeven wat er in de onderzoeksrapporten zelf wordt gelezen is de opdrachtformulering vaak cruciaal. Die bepaalt wat wel en niet wordt onderzocht. Onderzoekers zijn daarbij te veel geneigd de opdrachtformulering als heilig te beschouwen ook al leidt die tot beperkingen in het onderzoek die de conclusies twijfelachtig maken. Ook het gebruik van normenkaders lijkt gegeven de verscheidenheid en het ontbreken van een argumentatie waarom een bepaald normenkader is gebruikt, volstrekt willekeurig. Ten slotte nemen onderzoekers onwilligheid van betrokkenen om mee te werken aan het onderzoek maar al te vaak voor lief en trekken zij conclusies zonder aandacht te besteden aan dergelijke non-respons. Ook in de vertaling van de feiten naar de normatieve conclusies komen vaak drogredeneringen voor die de conclusie twijfelachtig maken.

Hoewel onderzoekers in hun weerwoord veel kritiek hebben geuit op de afsluiting die dit artikel al had voor het aan hun voor een weerwoord werd toegezonden, wil ik deze niet aan de lezer onthouden. Het gaat om de volgende uitsmijter, die dus niet de mening van de onderzoekers wier rapporten zijn beoordeeld weergeeft, maar wel van de schrijver van dit artikel, namelijk: De uitkomsten van dit

onderzoek doen denken aan een uitdrukking ontleend aan Multatuli, ‘Barbertje moet hangen’. Het ging om een man die beschuldigd werd van het vermoorden van zijn vrouw. Toen zijn vrouw levend de rechtszaal binnenkwam, moest de man alsnog hangen, nu vanwege zijn ‘eigenwaan’. Het is een uitdrukking geworden die aangeeft dat iemand schuldig moet worden verklaard, waarbij de vraag of dat terecht of onterecht is, geheel irrelevant is, en dat er altijd wel een reden is om iemand te veroordelen. Het lijkt er op dit bestuurders in de publieke sector ook kan overkomen en dat integriteitsonderzoek dit niet onaannemelijker maakt, maar juist faciliteert. Dat is betreurenswaardig.

8. Referenties

Addink, G. H. (2013). De toetsing van integriteitsschendingen getoetst: Een onderzoek naar het bestuurlijk-juridisch toetsingskader bij onderzoeken naar integriteitsschendingen in de publieke sector (pp. 1–25).

Bureau Integriteit Gemeente Amsterdam. (2006). Protocol integriteitonderzoeken gemeente Amsterdam.

Bureau Integriteit Nederlandse Gemeenten. (2012). Gemeente Wassenaar: Rapportage onderzoek inzake beschuldiging van ongewenst gedrag.

Bureau Integriteit Nederlandse Gemeenten. (2013). Rapportage onderzoek Feijenoord: Deelgemeente Feijenoord.

Bureau Integriteitsbevordering Openbare Sector. (2014). Handreiking onderzoek integriteits-schendingen politieke ambtsdragers.

Deloitte. (n.d.). Rapport inzake onderzoek naar de onkostenvergoedingen en declaraties van het college van Gedeputeerde Staten, (2010).

Dohmen, J. (2012, December 15). Commissie meldde niet alles over Van Rey. *NRC Handelsblad*.

Elzinga, D. J., & De Haan, E. (2015). Contra-expertise rapport “Bouwen aan de Barnseweg.”

Hoffmann Bedrijfsrecherche. (2015). Managementbrief II nr. 21406074: Betreft: Onderzoek i.o.v. Gemeente Urk te Urk.

KAFI Integrity. (2014). Uit Balans: Onderzoek naar enkele dossiers die behoren tot de Ruimtebalansprojecten in opdracht van de gemeente Nuenen.

KPMG. (2011). Gemeente Sliedrecht: Onderzoek relatie gemeente Sliedrecht en Baggerfestival Sliedrecht (Openbaar exemplaar).

NIVRA. Verordening gedragscode, registeraccountant, Pub. L. No. Artikel A-100.4.

Van den Heuvel, J. H. J. (2015a). Bouwen aan de Barnseweg.

Van den Heuvel, J. H. J. (2015b). Leren van feiten: Onderzoek naar de handelswijze van de Fractie Westland Verstandig tegenover de accountant.

Van den Heuvel, J. H. J., Pheijffer, M., & De Vries, J. (2012). *Ondernemend bestuur: Commissie Operatie Schoon Schip*.

Zouridis, S., & Vorm, B. Van Der. (2013). *Omwille van geloofwaardigheid: Een verkennend onderzoek naar knelpunten en oplossingen bij integriteitsonderzoek in Nederland*. Tilburg.